

A. Most strongly disease-associated patterns

Marker	01	02	03	04	05	06	07	08
Pattern01	4	1	6	6	3	*	*	*
Pattern02	4	1	6	6	3	6	*	*
Pattern03	4	1	6	6	*	6	*	*
Pattern04	4	1	6	*	3	6	*	*
Pattern05	4	1	*	6	3	*	*	*
Pattern06	4	1	*	6	3	6	*	*
Pattern07	4	*	6	6	3	*	*	*
Pattern08	4	*	6	6	3	6	*	*
Pattern09	4	*	6	*	3	6	*	*
Pattern10	*	1	6	6	*	*	*	*
Pattern11	*	1	6	6	3	*	*	*
Frequency	9	11	11	11	10	6	0	0

FIG. 1A

09875935.060804

103090" 5865/860

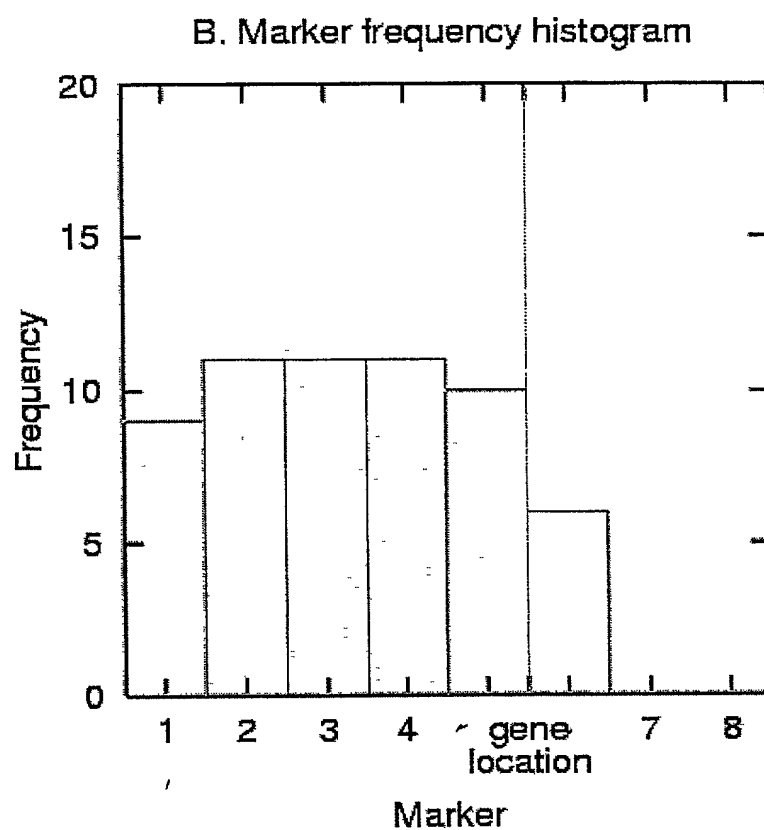


FIG 1B

C. Marker frequency histogram

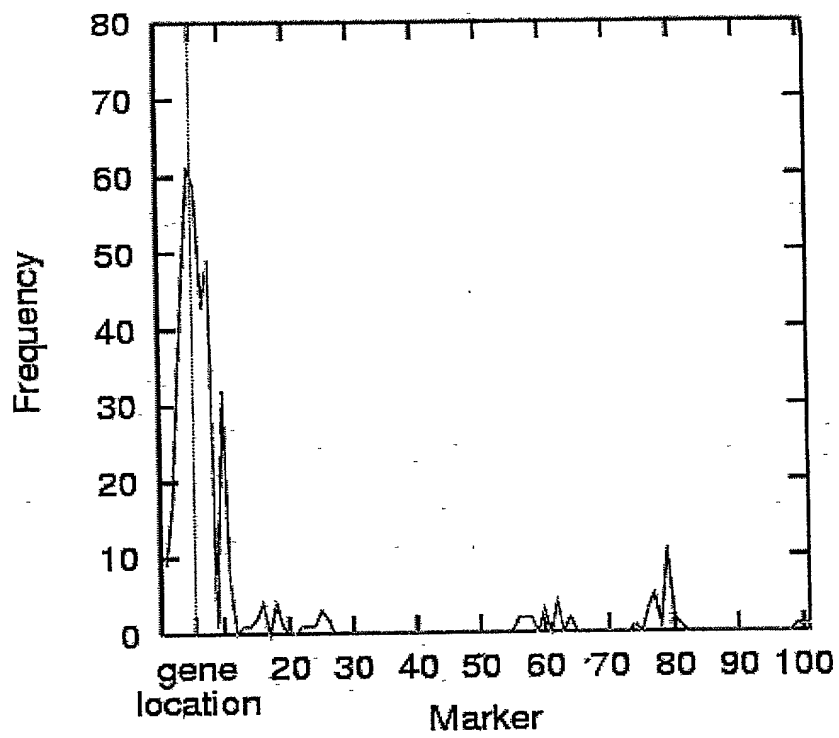


FIG. 1C

D. Prediction results (A = 10%)

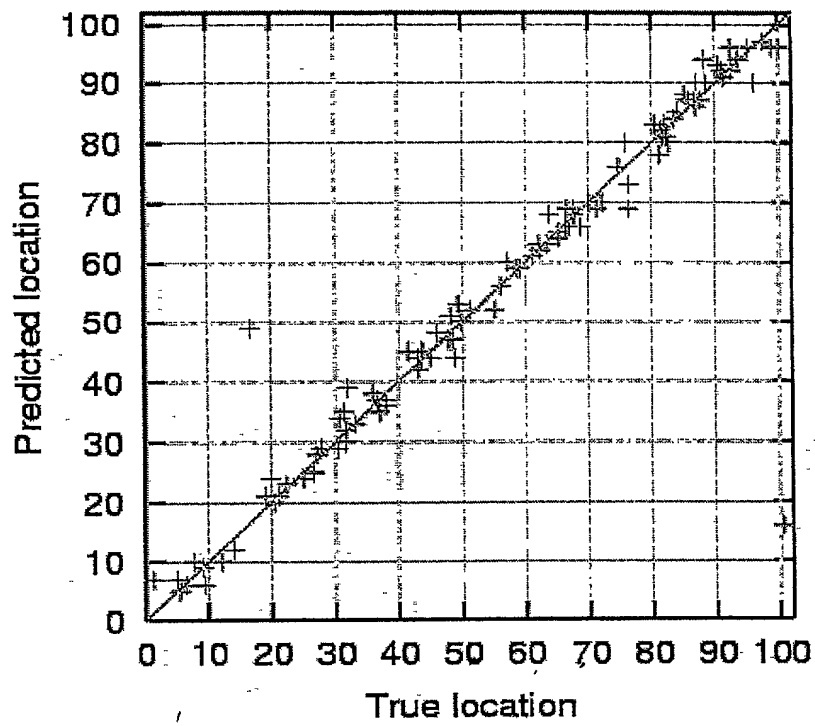


FIG. 1D

A. Influence of A

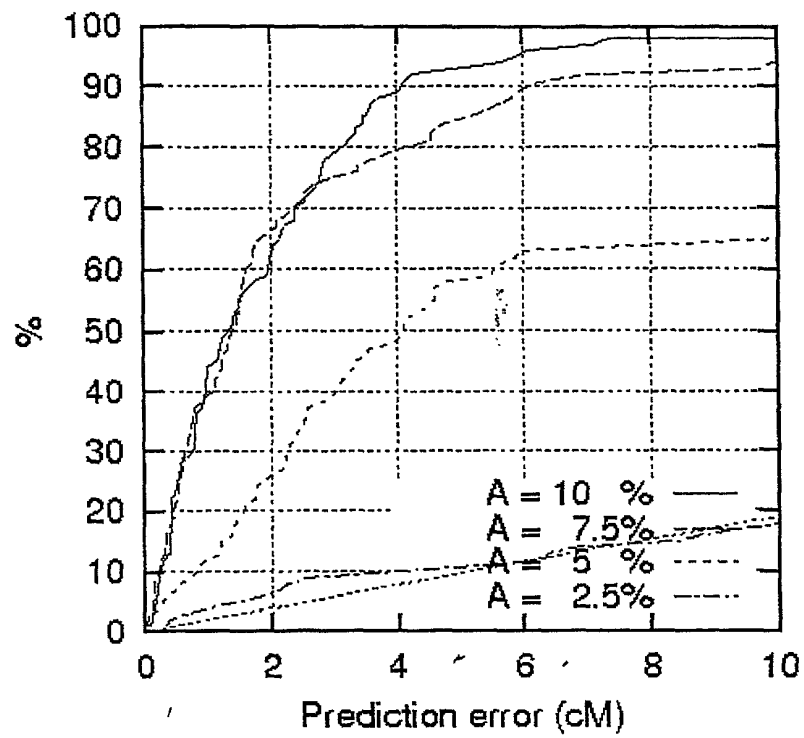


FIG. 2A

B. Influence of A (larger dataset)

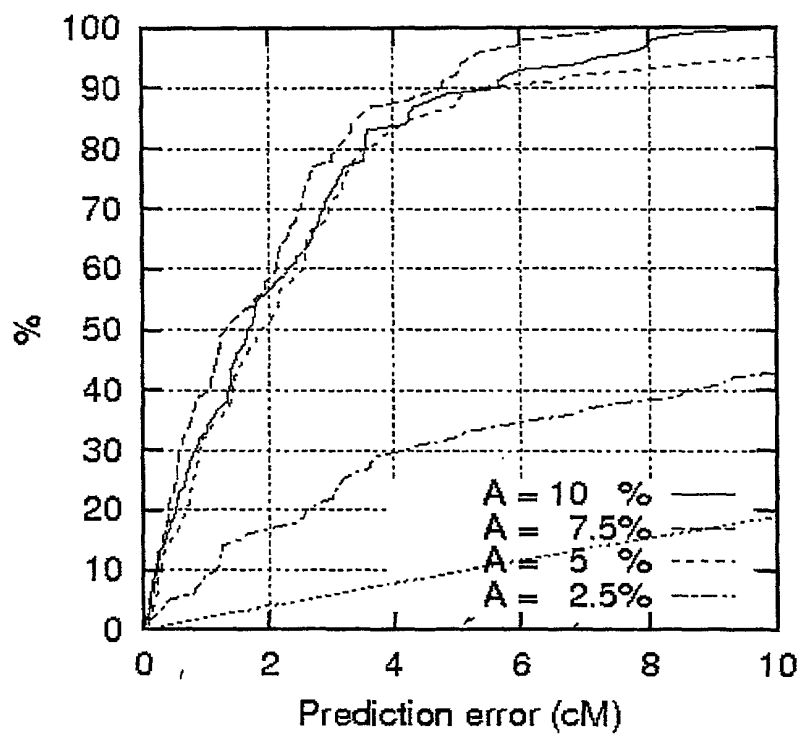


FIG. 2 B

C. Influence of corrupted data ($A = 10\%$)

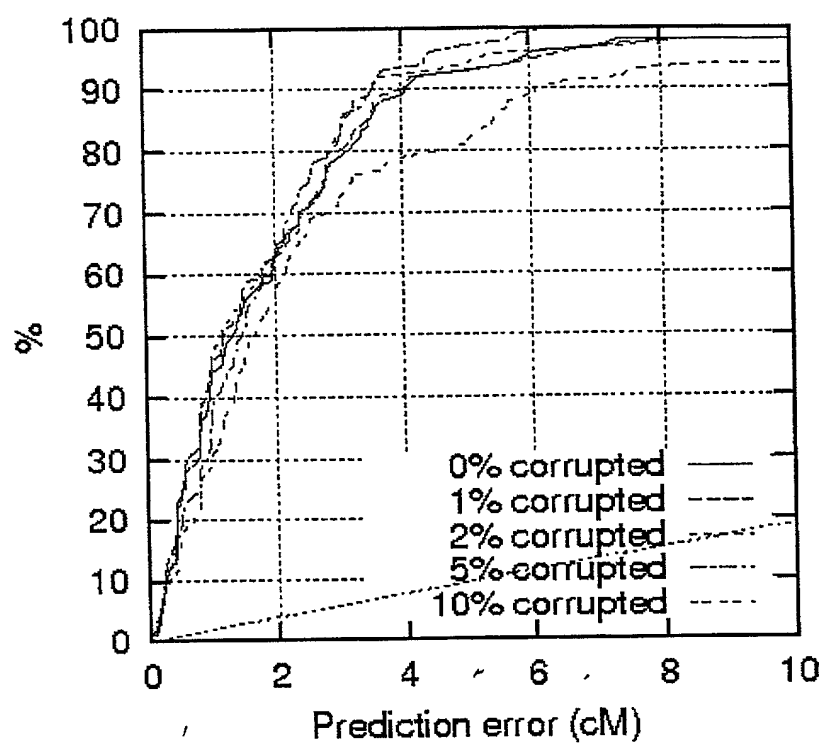


FIG. 2C

D. Influence of missing data (A = 10%)

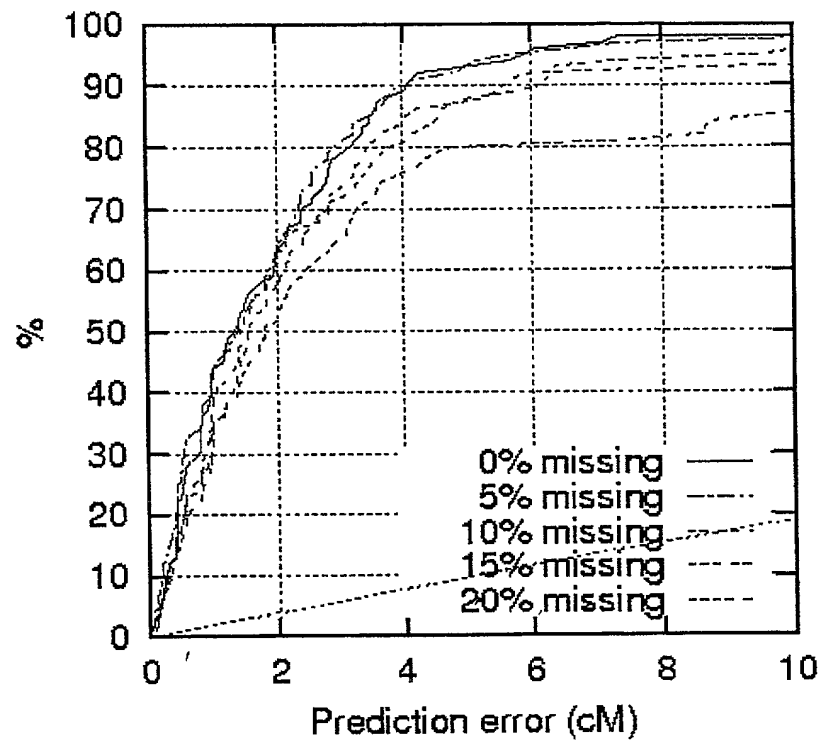


FIG. 2D

E. Comparison of prediction methods
(A = 10%, 0/1% corrupted, 0/20% missing)

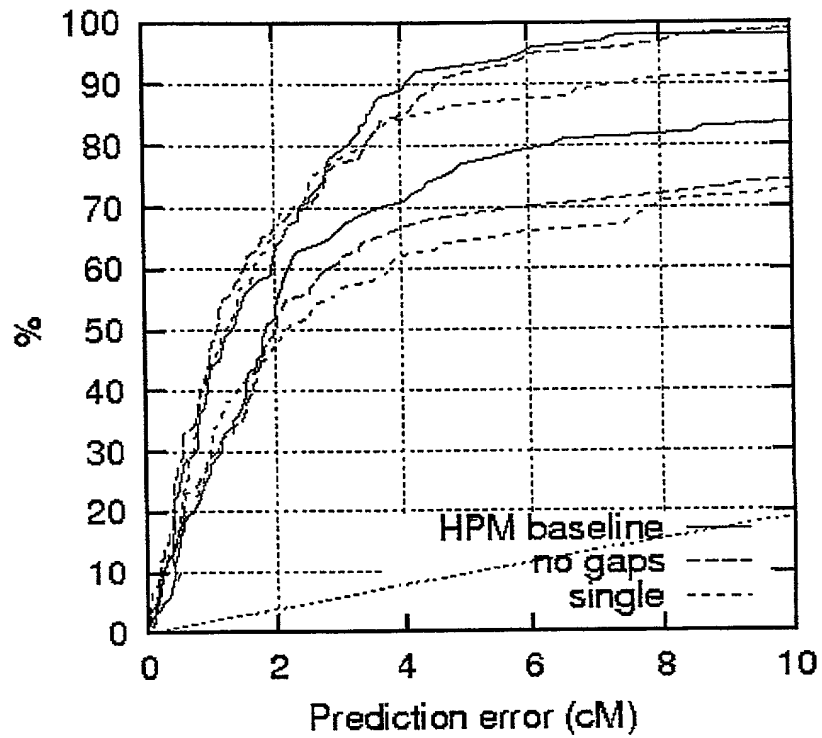


FIG. 2E

F. Influence of pattern parameters
(A = 10%, 1% corrupted, 20% missing)

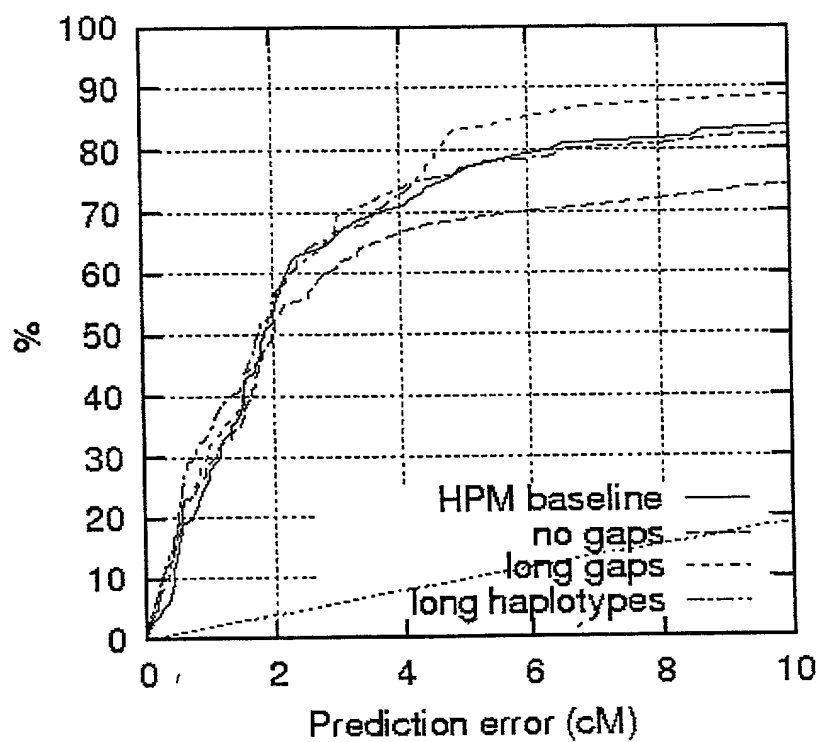


FIG. 2F

09875935.1060801
108090" 5E65/860

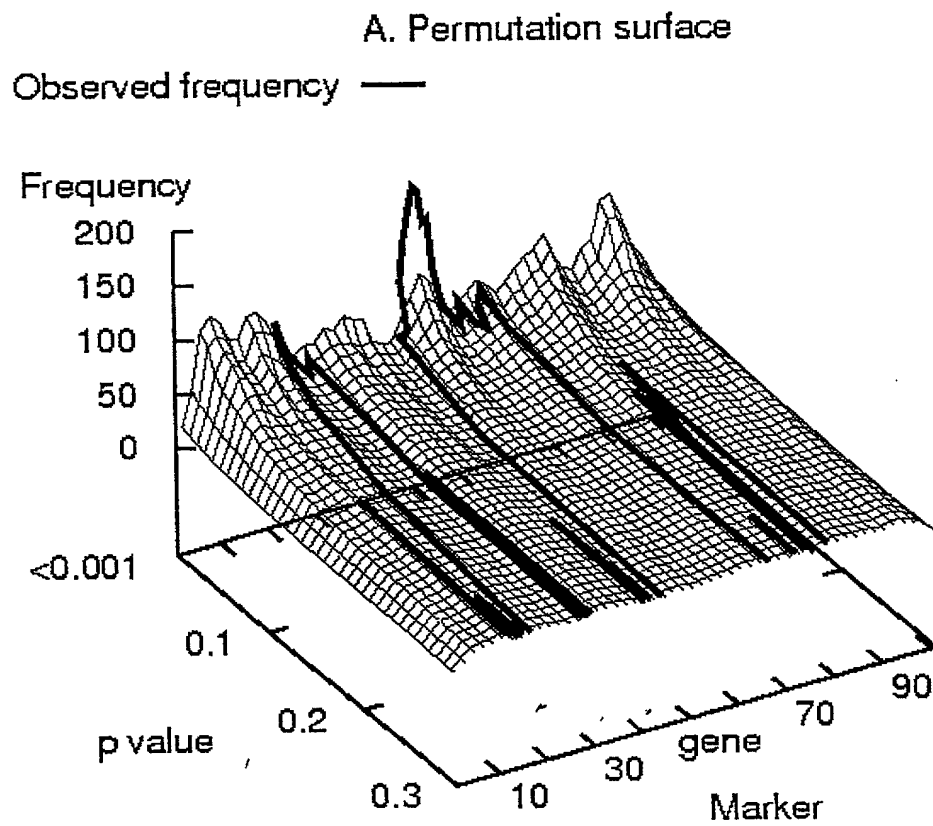


FIG. 3A

B. Marker-wise frequencies for some p values

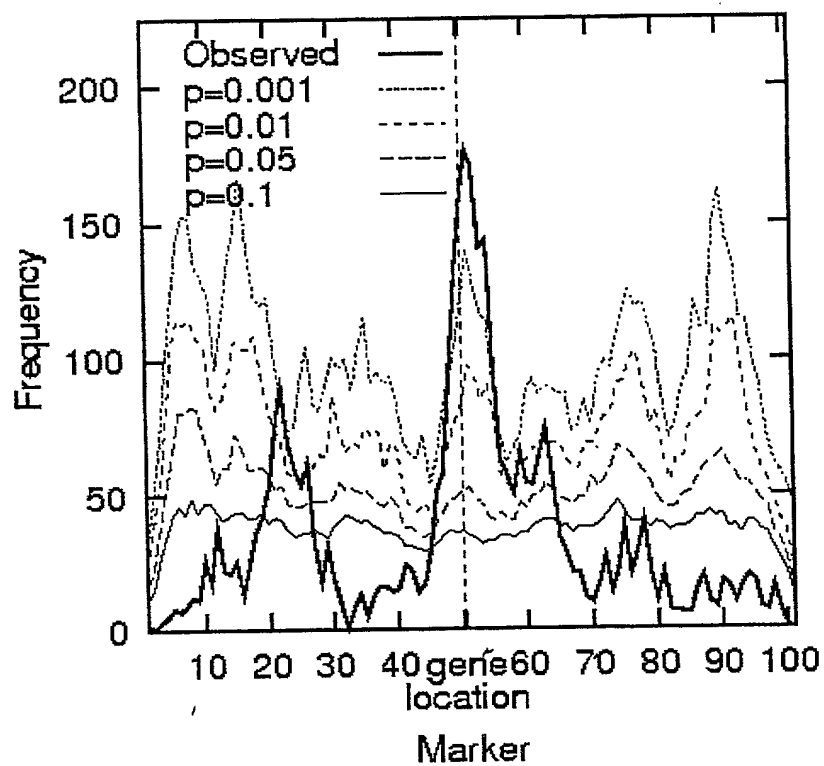


FIG. 3B

C. Marker-wise frequencies for some p values

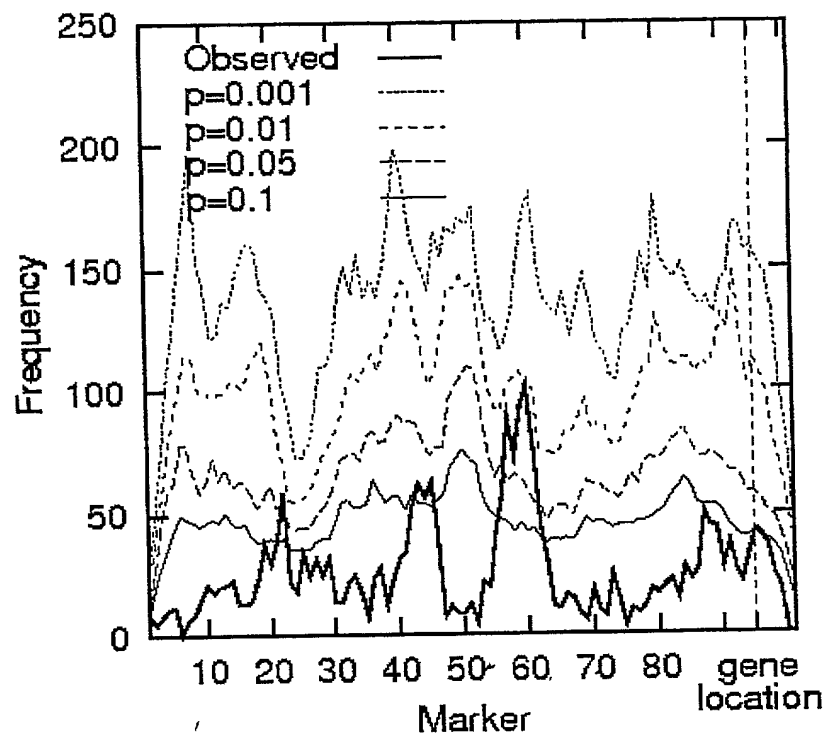


FIG. 3C

D. Influence of permutation tests ($A = 5\%$)

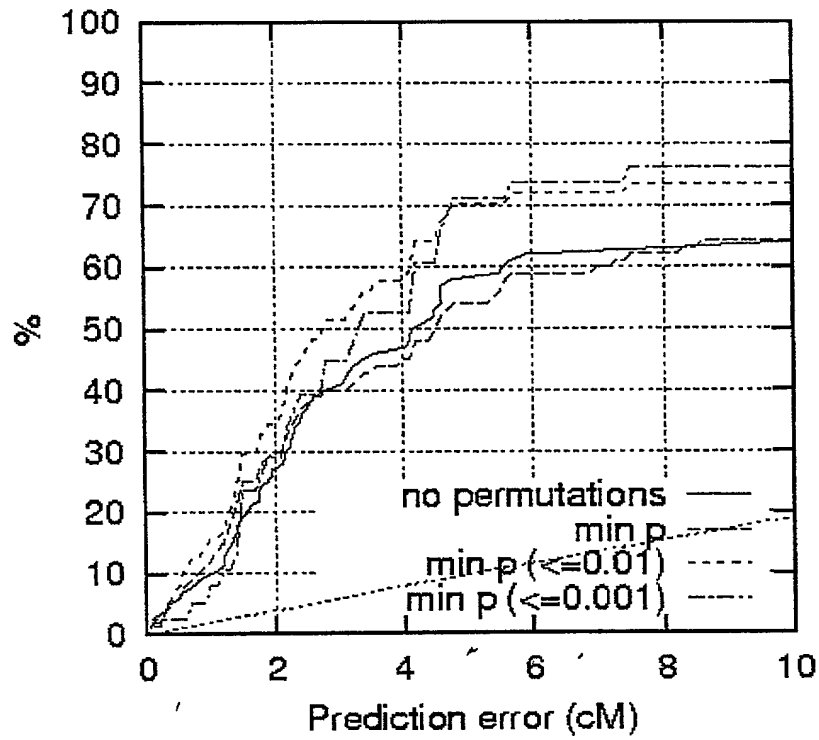


FIG. 3D

T08090" 5E654860

Influence of A (SNP dataset, 12.5% missing data)

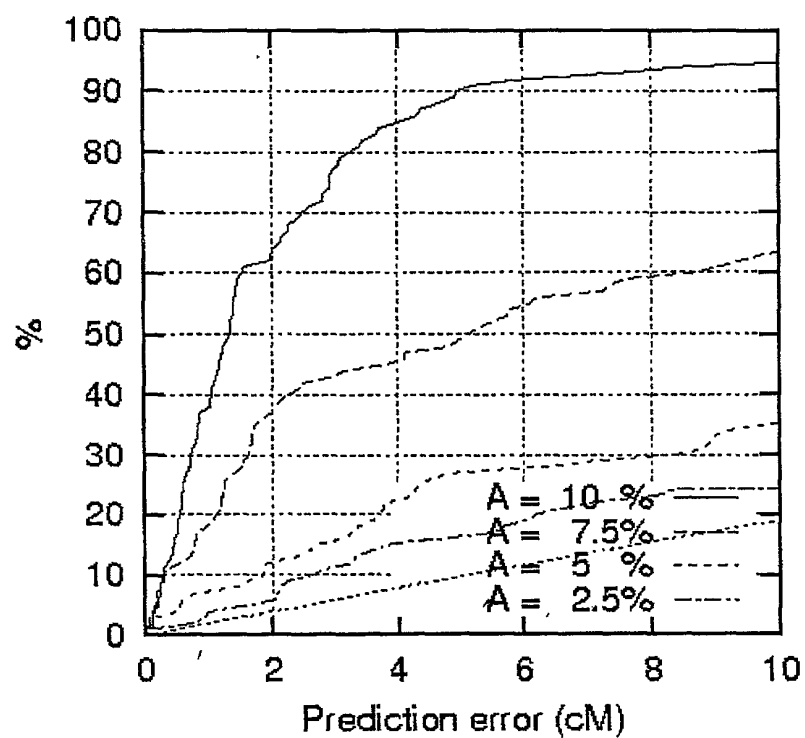


FIG. 4

A. Marker frequencies in HLA data

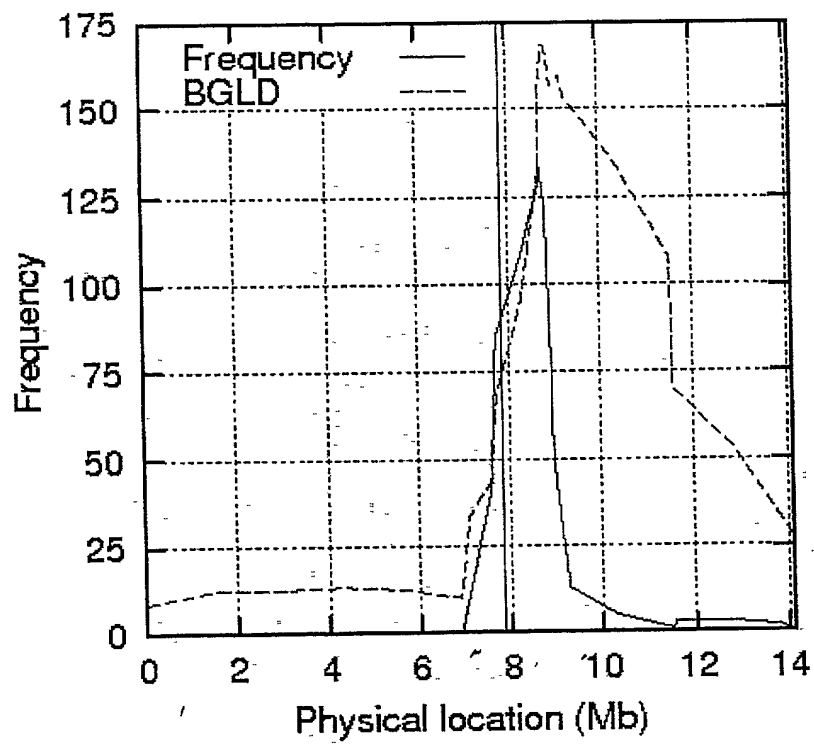


FIG. 5A

B. P values and frequency ratios in HLA data

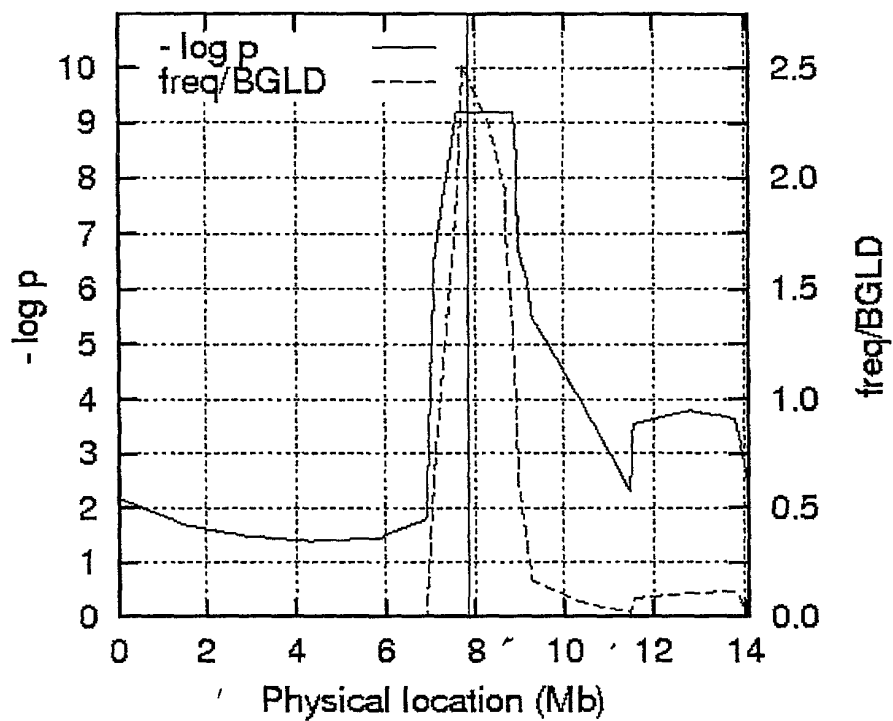


FIG. 5B

A. Marker frequency histogram (2 genes)

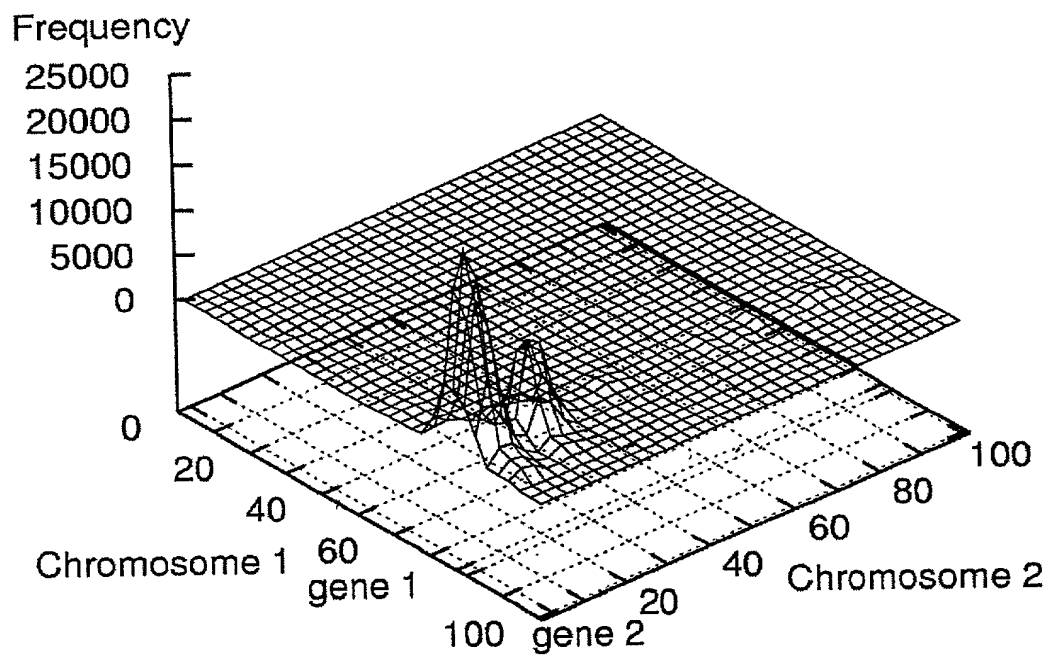


FIG. 6A

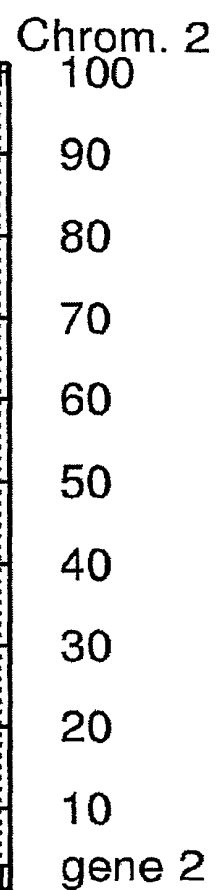
[illegible]

FIG. 6B

C. Localization residuals (2 genes)

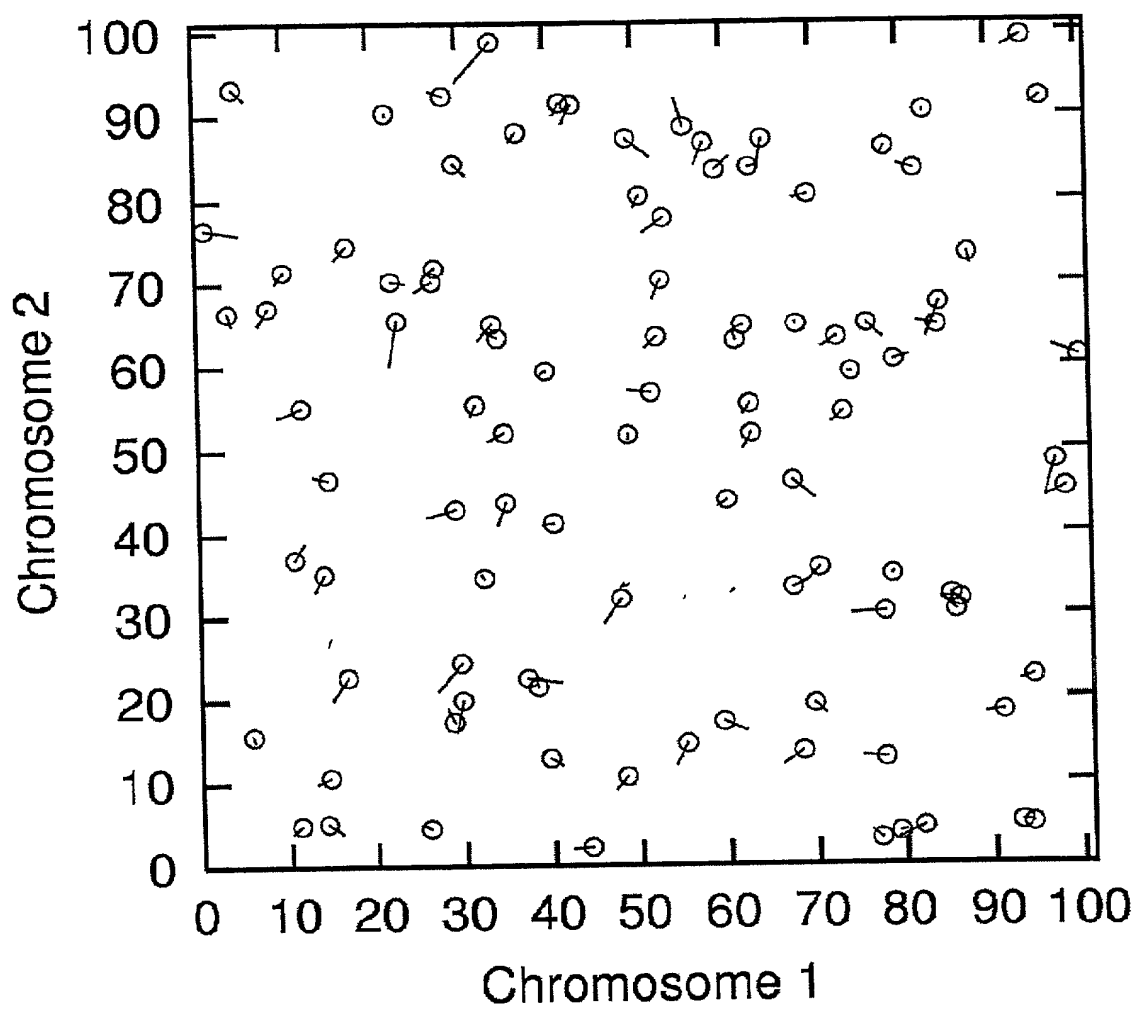


FIG. 6C

D. Prediction accuracy (2 genes)

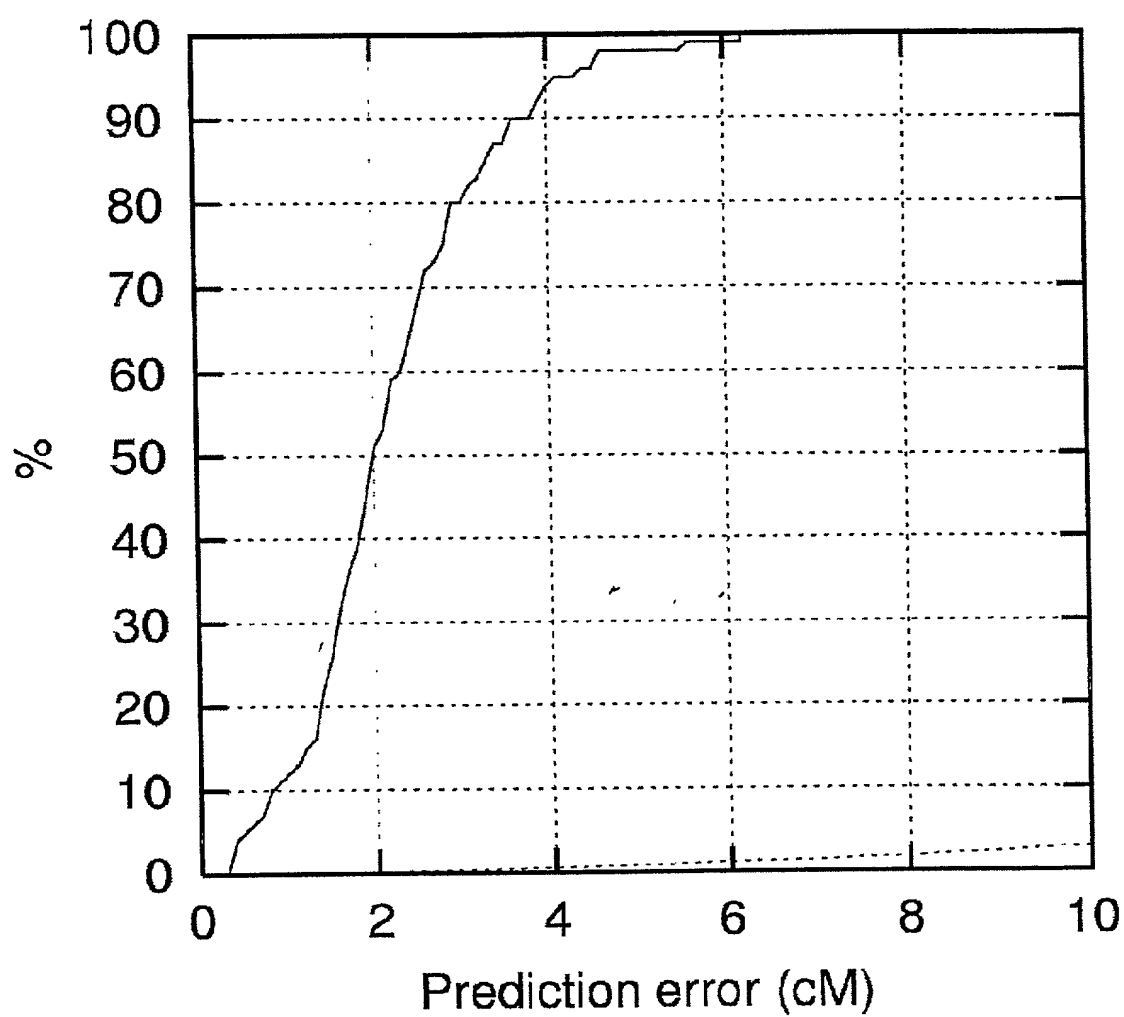


FIG. 6D

103030" 5654360

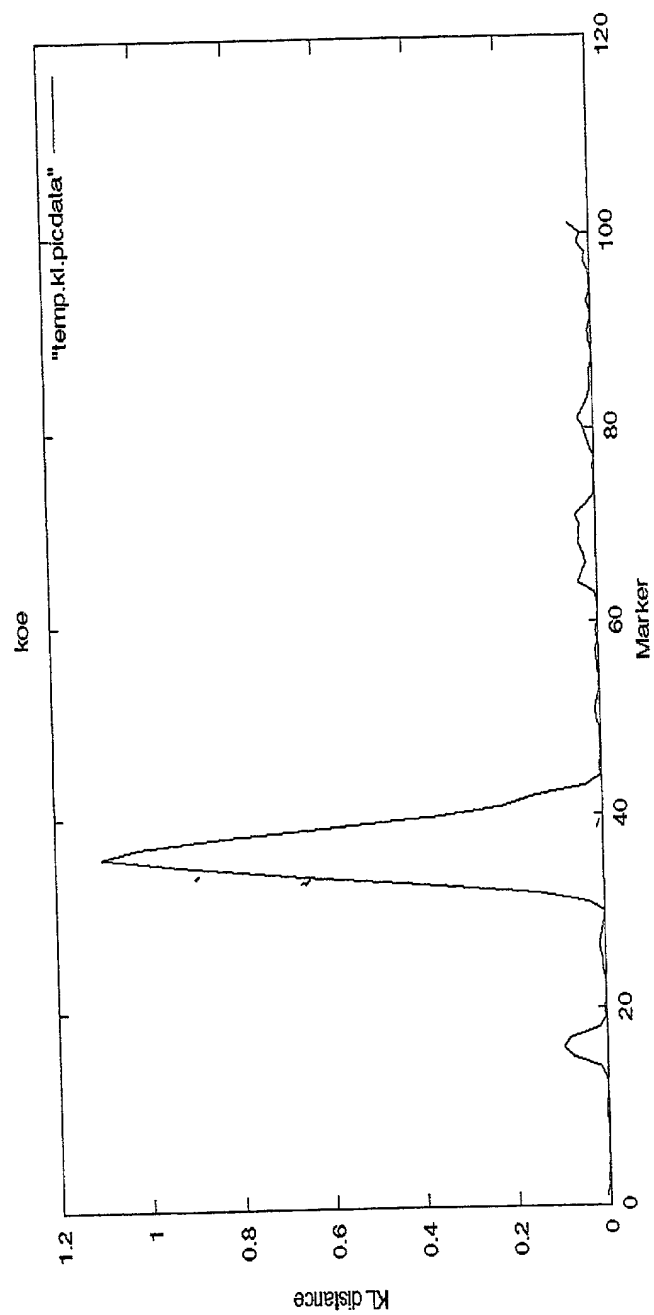


FIG. 7A

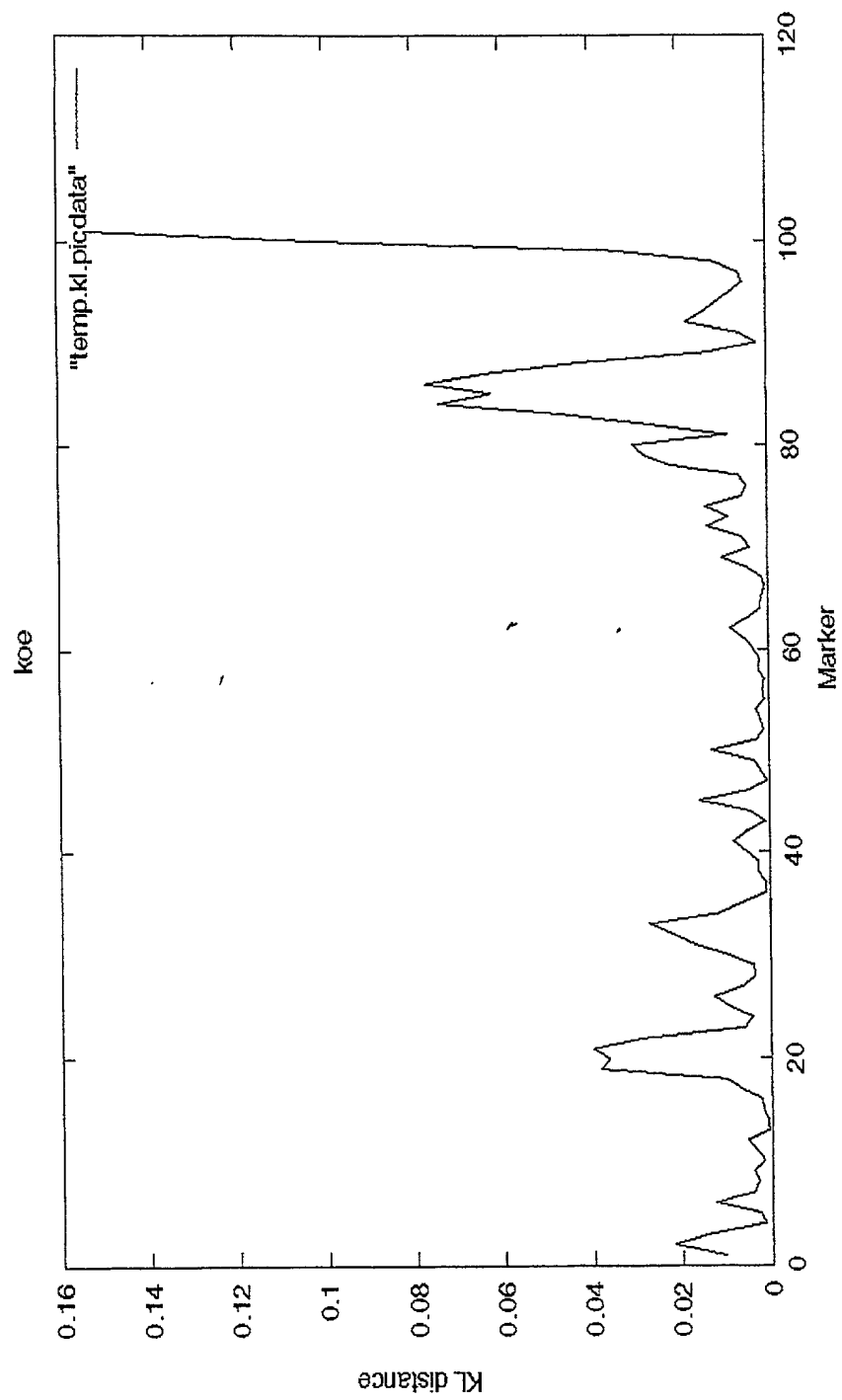


FIG. 7B